Available online at www.sciencedirect.com
ScienceDirect
Automatica 43 (2007) 1519–1531
automatica
www.elsevier.com/locate/automatica

Development of omni-directional correlation functions for nonlinear model validation☆

Quan Min Zhu[a,b,*], Li Feng Zhang[a], Ashley Longden[a]

[a]Faculty of Computing, Engineering and Mathematical Sciences, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK
[b]School of Communication and Control Engineering, Southern Yangtze University, Huihe Road 170, Wuxi 214036, China

Received 2 November 2005; received in revised form 6 November 2006; accepted 15 February 2007

Abstract

In the present study a set of first order correlation functions are proposed to examine the quality of a wide class of identified nonlinear models. The first order correlation functions, defined as omni-directional correlation functions, are integrated into two concise tests to provide more effective auto and cross model error correlation diagnosis than the other approaches proposed from higher order correlation functions. The mechanisms of the novel validity tests are proved in theory and demonstrated with numerical analyses. Two simulated case studies, in the situation of incorrectly detected model structure and estimated parameters, are presented to illustrate the diagnostic power of the new methodology.

Keywords: Correlation; Model validation; Nonlinear systems; Dynamical systems; System identification

1. Introduction

Nowadays, nonlinear dynamic modelling is applied in many fields to describe a wide range of natural and manmade systems. Model validation is the final step of any modelling procedure to check the quality of the identified model, and to determine if the model is an appropriate representative of the underlying system. Developing concise and efficient nonlinear model validity test methodologies has been a challenging research and received wide attention for over a half century (Bohlin, 1971; Box & Jenkins, 1976; Carson & Flood, 1990; Dullerud & Smith, 2002; Parke, Holford, & Charles, 1999; Rangan & Poolla, 1996). Although most linear model validation test approaches cannot be directly used to examine identified nonlinear model fitness, they do provide useful references for nonlinear model validation. Therefore, it is pertinent at the beginning of the study to explain linear model validation context, solutions, and limitations.

If a system under analysis is linear, a number of well-established methodologies can be used for validating the identified model. Some of the most powerful methods are based on the concept that if the model structure is correct and the parameters estimation is unbiased, the residuals should form a random sequence with zero mean and finite variance. Auto-correlation function (ACF) and cross-correlation function (CCF), therefore, have been widely applied in linear model validation and the studies of Bohlin (1971, 1978), Box and Jenkins (1976), and Soderstrom and Stoica (1990) show that the ACF of residuals and the CCF between residuals and inputs should lie within a preset confidence interval when the identified model is correct and the residual sequence is completely random.

Unfortunately, nonlinear model validation is not as straightforward as linear model validation. Bohlin's ACF and CCF tests are obviously inadequate for validating nonlinear models since they cannot diagnose all possible missing nonlinear terms in the residuals (Billings & Voon, 1983). To validate nonlinear models, several correlation tests based approaches have been

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Johan Schoukens under the direction of Editor Torsten Söderström.

* Corresponding author. Faculty of Computing, Engineering and Mathematical Sciences, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK. Tel.: +44 117 32 82533; fax: +44 117 32 83800.

E-mail address: quan.zhu@uwe.ac.uk (Q.M. Zhu).

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2007.02.010

developed such as higher order CCF tests (Aguirre, 1995, 1997; Billings & Voon, 1983; Hjalmarsson, 1993; Zhu & Billings, 1997), a combination of five first and second order ACF and CCF tests (Billings & Voon, 1986), higher order ACF and CCF between outputs, inputs and residuals (Billings & Zhu, 1994, 1995), and multi-directional correlation tests (Mao & Billings, 2000). These methods have been successfully applied in practical system identification (Aguirre & Billings, 1994; Billings, Chen, & Backhouse, 1989; Sriniwas, Arkun, Chien, & Ogunnaike, 1995; Thomson, Schooling, & Soufian, 1996). Nevertheless, these approaches do not perform adequately when some special nonlinear effects such as even power terms occur in the residuals (Zhang, Zhu, & Longden, 2007). This means that the correlation functions, under certain conditions, may still fall inside the 95% confidence interval even if predictable components are present in the residuals. In addition, higher order correlation functions can sometimes exhibit less power when the variances of noise and input are small, because of the fourth and higher moments become small (Billings & Voon, 1986).

To deal with these problems, this study is devoted to developing a new approach based on a set of first order correlation functions. A significant progression to enhance the detection power is to classify model nonlinear terms into four different types of nonlinear associations. Accordingly the detecting power of the first order correlation functions is retained while no nonlinear terms are missed out from the detecting net. The remainder of the study is organised as follows. In Section 2 omni-directional cross-correlation functions (ODCCFs) are proposed to weave a correlation detecting net to find out all linear and nonlinear associations in a regression model. In Section 3 these proposed ODCCFs are integrated into a concise and efficient formulation set, called combined ODCCF, to test the autocorrelation and cross correlation of model residuals and inputs. In Section 4 two improperly identified nonlinear models are selected to illustrate the detection power using the new model validity tests. Finally in Section 5 conclusions are drawn to summarise the study. To reduce the length of the study and to facilitate readability, all the relevant theoretical proofs and numerical demonstrations are enclosed in appendices.

2. Omni-directional correlation functions (ODCCFs)

Consider a relationship between dependent variable $\mathbf{y}$ and independent variable $\mathbf{x}$ as below.

$$y(n) = f(\mathbf{x}^{n-1}, \mathbf{z}^{n-1}), \tag{1}$$

where n ($n = 1, 2, \ldots, N$) is a time index, $f(\cdot)$ is a linear or nonlinear function, and

$$\left.\begin{aligned} \mathbf{x}^{n-1} &= [x(n-1), \ldots, x(n-r)], \\ \mathbf{z}^{n-1} &= [z(n-1), \ldots, z(n-r)] \end{aligned}\right\} \tag{2}$$

with delayed elements from 1 to r.

Remark 1. It should be noticed that $\mathbf{z}$ denotes the dummy variable which represents the effects from all the other factors except $\mathbf{x}$. For a noise free SISO system, $\mathbf{z}$ can be considered as a zero vector. For a SISO system which is submerged in a noisy environment, $\mathbf{z}$ can be considered as an additive Gaussian white noise. For a system with multi-inputs, $\mathbf{z}$ can be considered as including all the other independent variables except $\mathbf{x}$.

To describe the relationships between $y(n)$ and $x(n-\tau)$, two definitions are set up as below (Zhang et al., 2007).

Definition 1. Symmetry along dependent variable axis: A constant time k exists that for any time m, there is a time p so that

$$f(\mathbf{x}^{n-m}, \mathbf{z}^{n-m}) = f((\mathbf{x}^{n-p})^*, \mathbf{z}^{n-p})$$

where

$$\begin{aligned}(\mathbf{x}^{n-p})^* = [&x(n-p), \ldots, x(n-p-\tau+1), 2x(k) \\ &- x(n-m-\tau), x(n-p-\tau-1), \\ &\ldots, x(n-p-r)]\end{aligned}$$

and

$$x(k) \neq x(n-m-\tau). \tag{3}$$

Definition 2. Symmetry along independent variable axis: A constant h exists that for any time m, there is a time q so that

$$\begin{aligned} f(\mathbf{x}^{n-m}, \mathbf{z}^{n-m}) - f(\mathbf{x}^{n-h}, \mathbf{z}^{n-h}) = f(\mathbf{x}^{n-h}, \mathbf{z}^{n-h}) \\ - f((\mathbf{x}^{n-q})^*, \mathbf{z}^{n-q}), \end{aligned}$$

where

$$\begin{aligned}(\mathbf{x}^{n-q})^* = [&x(n-q), \ldots, x(n-q-\tau+1), x(n-m-\tau), \\ &x(n-q-\tau-1), \ldots, x(n-q-r)]\end{aligned}$$

and

$$f(\mathbf{x}^{n-m}, \mathbf{z}^{n-m}) \neq 0. \tag{4}$$

The two symmetrical properties are shown in Figs. 1 and 2 by using two simple examples with $x(k)=0$ and $f(\mathbf{x}^{n-h}, \mathbf{z}^{n-h})=0$, respectively.

With the above definitions, all nonlinear associations between any two variables in (1) can be classified into four categories, which cover all the possible nonlinear effects concerned with both the amplitude and the sign of each variable. For describing the relationship between two data sequences with a specified delay time, $y(n)$ and $x(n-\tau)$ are considered as a pair of individual dependent variable and independent variable. After removing the mean level from each data sequence, the relationship between $y(n)$ and $x(n-\tau)$ are classified as follows.

Type 1: The amplitude of the dependent variable varies as the amplitude of the independent variable varies, thus both symmetrical properties are satisfied.

Type 2: Both the sign and amplitude of the dependent variable vary as the amplitude of the independent variable varies, thus only the first symmetrical property is satisfied.

Type 3: Both the sign and amplitude of the dependent variable vary as both the sign and the amplitude of the independent

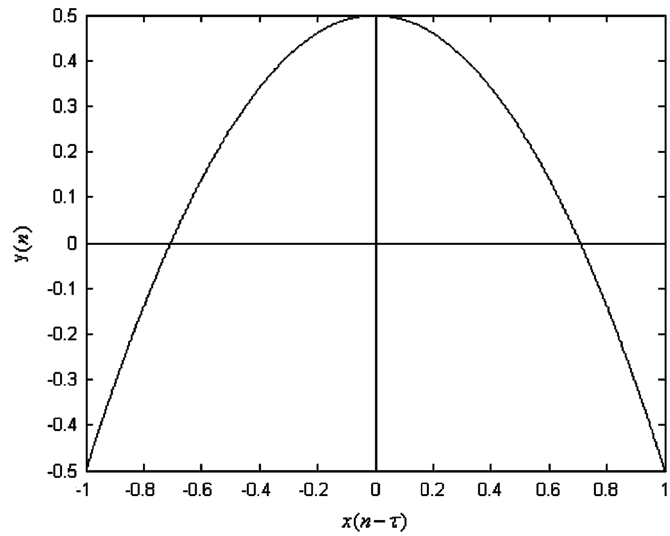


Fig. 1. The shape of scatter plot symmetry along dependent variable axis.

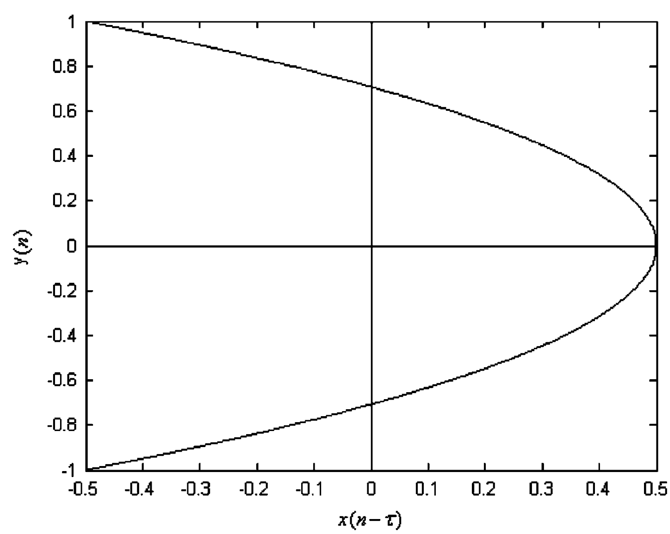


Fig. 2. The shape of scatter plot symmetry along independent variable axis.

variable vary, thus both symmetrical properties are not satisfied. (Linear association is a special case of this type of nonlinear association.)

Type 4: The amplitude of the dependent variable varies as both the sign and the amplitude of the independent variable vary, thus only the second symmetrical property is satisfied.

To illustrate such classifications, three simple examples are shown as follows. For simplicity but not losing generality, there is no delayed term included in the models and all the inputs are uniformly distributed uncorrelated data sequences with zero mean and finite variance.

$$\begin{cases} y_1(n) = x_1(n)x_2(n), \\ y_2(n) = x_1^2(n), \\ y_3(n) = x_2(n)(x_1(n) + a). \end{cases} \tag{5}$$

In (5), a is a constant defined as

$$a \geqslant \max(|\mathbf{x}_1|). \tag{6}$$

Consider the first model in (5), the relationships between $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{y}_1$ exhibit the first type of nonlinear association and satisfy both symmetrical properties. For $\mathbf{x}_1$, a constant time k_1 ($x_1(k_1) = 0$) exists that for any time m_1 there is a time p_1 ($x_2(n - p_1) = -x_2(n - m_1)$) so that $x_1(n - m_1)x_2(n - m_1) = (2x_1(k_1) - x_1(n - m_1))x_2(n - p_1)$. And a constant time h_1 ($y_1(h_1) = 0$) exists that for any time m_2 there is a time q_1 ($x_2(n - q_1) = -x_2(n - m_2)$) so that $x_1(n - m_2)x_2(n - m_2) - y_1(h_1) = y_1(h_1) - x_1(n - m_2)x_2(n - q_1)$. For $\mathbf{x}_2$, a similar analysis can be used.

For the second model, the relationship between $\mathbf{x}_1$ and $\mathbf{y}_2$ exhibits the second type of nonlinear association and satisfies the first symmetrical property. A constant time k_2 ($x_1(k_2) = 0$) exists that for any time m_3 that is $x_1^2(n - m_3) = (2x_1(k_2) - x_1(n - m_3))^2$.

For the third model in (5), the relationship between $\mathbf{x}_1$ and $\mathbf{y}_3$ exhibits the fourth types of nonlinear association and satisfies the second symmetrical property. A constant time h_2 ($y_3(h_2)$=0) exists that for any time m_4 there is a time q_2 ($x_2(n - q_2) = -x_2(n - m_4)$) so that $x_2(n - m_4)(x_1(n - m_4) - a) - y_3(h_2) = y_3(h_2) - x_2(n - q_2)(x_1(n - m_4) - a)$. The relationship between $\mathbf{x}_2$ and $\mathbf{y}_3$ exhibits the third type of nonlinear association since it does not satisfy both symmetrical properties.

Remark 2. Based on definitions 1 and 2, four types of nonlinear associations can be classified. That is, the definitions provide rules to classify all possible nonlinear associations concerned with both the amplitude and the sign of each variable into four groups (types).

Remark 3. While detecting a correlationship between dependent variable $\mathbf{y}$ and independent variable $\mathbf{x}$ with a specified delay time (such as $y(n) = x^3(n - \tau)$, $\tau = 5$). The relationship can be considered as static at $\tau = 5$. The dynamic relationship between $\mathbf{y}$ and $\mathbf{x}$ is the whole process with all interested delay times ($\tau = 0, 1, \ldots, 10$, say).

Consequently, a set of normalised first order correlation functions, named as omni-directional cross-correlation functions (ODCCFs), are proposed to detect above these four types of nonlinear associations.

To detect type 1 associations

$$r_{\beta'\alpha'}(\tau) = \frac{\sum_{n=\tau+1}^{N} \alpha'(n)\beta'(n - \tau)}{[(\sum_{n=1}^{N} (\alpha'(n))^2)(\sum_{n=1}^{N} (\beta'(n))^2)]^{1/2}}. \tag{7}$$

To detect type 2 associations

$$r_{\beta'y'}(\tau) = \frac{\sum_{n=\tau+1}^{N} y'(n)\beta'(n - \tau)}{[(\sum_{n=1}^{N} (y'(n))^2)(\sum_{n=1}^{N} (\beta'(n))^2)]^{1/2}}. \tag{8}$$

To detect type 3 associations

$$r_{x'y'}(\tau) = \frac{\sum_{n=\tau+1}^{N} y'(n)x'(n - \tau)}{[(\sum_{n=1}^{N} (y'(n))^2)(\sum_{n=1}^{N} (x'(n))^2)]^{1/2}}. \tag{9}$$

To detect type 4 associations

$$r_{x'\alpha'}(\tau) = \frac{\sum_{n=\tau+1}^{N}\alpha'(n)x'(n-\tau)}{[(\sum_{n=1}^{N}(\alpha'(n))^2)(\sum_{n=1}^{N}(x'(n))^2)]^{1/2}}, \quad (10)$$

where the prime $'$ in (7)–(10) denotes that the mean level has been removed from the corresponding data sequence, and

$$\left.\begin{aligned} \alpha(n) &= |y'(n)| = \left|y(n) - \frac{1}{N}\sum_{1}^{N} y(n)\right|, \\ \beta(n) &= |x'(n)| = \left|x(n) - \frac{1}{N}\sum_{1}^{N} x(n)\right| \end{aligned}\right\} \quad (11)$$

and are normalised via

$$\left.\begin{aligned} \alpha'(n) &= \alpha(n) - \frac{1}{N}\sum_{1}^{N}\alpha(n), \\ \beta'(n) &= \beta(n) - \frac{1}{N}\sum_{1}^{N}\beta(n). \end{aligned}\right\} \quad (12)$$

See appendices for the analytical proofs and numerical demonstrations of the (5) that (7)–(10) can be used to correspondingly detect the four types of nonlinear associations.

For a special case $\mathbf{x} = \mathbf{y}$, the functions (7)–(10) are called omni-directional auto-correlation functions (ODACFs). For whatever nonlinear term exists in the relationship of (1), there are one or more functions that can properly detect its correlation with the dependent variable. In addition, these functions avoid a disadvantage of higher order correlation tests which is that they can sometimes exhibit less detection power when the variances of input and noise are small to make which in consequence the fourth and higher moments become very small (Billings & Voon, 1986).

3. Combined ODCCF and ODACF for model validation

The mechanism of the new correlation functions has been efficiently applied to detect nonlinear relationships (Zhang et al., 2007). In this study the correlation functions are integrated to form a concise nonlinear model validity test procedure. Consider a generalised single input and single output (SISO) nonlinear parametric model.

$$y(n) = \hat{y}(n) + \varepsilon(n) = \hat{f}(\mathbf{y}^{n-1}, \mathbf{u}^{n-1}, \boldsymbol{\epsilon}^{n-1}) + \varepsilon(n), \quad (13)$$

where $\hat{f}(\cdot)$ is the identified nonlinear model. $\mathbf{y}^{n-1}$, $\mathbf{u}^{n-1}$, and $\boldsymbol{\epsilon}^{n-1}$ are measured output, input and residual (the difference between the measured output and the model predictive output) vectors with delayed elements from 1 to r, respectively. It should be noticed that when above model is properly identified, residual $\varepsilon(n)$ should be reduced to a random noise sequence denoted by $e(n)$ with zero mean and finite variance (Billings & Zhu, 1994). In other words, a residual sequence which is obtained from using a valid identified model should be uncorrelated to the delayed residuals, inputs and outputs (Ljung, 1999).

It has been analysed (Zhang et al., 2007) that the proposed correlation tests can be combined in a more concise and efficient prototype to examine nonlinear associations. In principle this can be expanded for the examination of validity of identified nonlinear models. To develop the combined formulations, first of all a set of omni-directional auto-correlation functions (ODACFs) of residuals and omni-directional cross-correlation functions (ODCCFs) between residuals and delayed inputs are proposed as follows.

When a model is properly identified, its residual should be reduced to an uncorrelated zero mean noise that is $\varepsilon(n) = e(n)$, ODACFs for residuals are formulated as

$$\begin{cases} r_{\alpha'\alpha'}(\tau) = \begin{cases} 1, & \tau = 0, \\ 0 & \text{otherwise}, \end{cases} \\ r_{\alpha'\varepsilon'}(\tau) = 0 \quad \forall\tau, \\ r_{\varepsilon'\varepsilon'}(\tau) = \begin{cases} 1, & \tau = 0, \\ 0 & \text{otherwise}, \end{cases} \\ r_{\varepsilon'\alpha'}(\tau) = 0 \quad \forall\tau. \end{cases} \quad (14)$$

ODCCFs between residuals and delayed inputs are formulated as

$$\begin{cases} r_{\beta'\alpha'}(\tau) = 0 \quad \forall\,\tau, \\ r_{\beta'\varepsilon'}(\tau) = 0 \quad \forall\,\tau, \\ r_{u'\varepsilon'}(\tau) = 0 \quad \forall\,\tau, \\ r_{u'\alpha'}(\tau) = 0 \quad \forall\,\tau, \end{cases} \quad (15)$$

where

$$\left.\begin{aligned} \alpha'(n) &= |\varepsilon'(n)| - \frac{1}{N}\sum_{1}^{N}|\varepsilon'(n)|, \\ \beta'(n) &= |u'(n)| - \frac{1}{N}\sum_{1}^{N}|u'(n)|. \end{aligned}\right\} \quad (16)$$

The magnitudes of ODACFs and ODCCFs are comparable since they are all calculated based on the first order correlation principles, which means that a higher value of ODACFs or ODCCFs denotes a more significant nonlinear association whichever a function reports the correlation. The results obtained from ODACFs and ODCCFs, therefore, can be concisely and efficiently integrated into two new correlation functions named combined ODACF ($\rho_{\varepsilon\varepsilon}(\tau)$) and combined ODCCF ($\rho_{u\varepsilon}(\tau)$). They are defined as follows.

Firstly, let

$$\begin{cases} R_{\varepsilon\varepsilon}(\tau) = [r_{\alpha'\alpha'}(\tau), r_{\alpha'\varepsilon'}(\tau), r_{\varepsilon'\varepsilon'}(\tau), r_{\varepsilon'\alpha'}(\tau)], \\ R_{u\varepsilon}(\tau) = [r_{\beta'\alpha'}(\tau), r_{\beta'\varepsilon'}(\tau), r_{u'\varepsilon'}(\tau), r_{u'\alpha'}(\tau)]. \end{cases} \quad (17)$$

Definition 3. Combined ODACF

$$\begin{cases} \text{if}, |\max(R_{\varepsilon\varepsilon}(\tau))| > |\min(R_{\varepsilon\varepsilon}(\tau))|, \\ \rho_{\varepsilon\varepsilon}(\tau) = \max(R_{\varepsilon\varepsilon}(\tau)), \\ \text{else} \\ \rho_{\varepsilon\varepsilon}(\tau) = \min(R_{\varepsilon\varepsilon}(\tau)). \end{cases} \quad (18)$$

Definition 4. Combined ODCCF

$$\begin{cases} \text{if}, |\max(R_{u\varepsilon}(\tau))| > |\min(R_{u\varepsilon}(\tau))|, \\ \rho_{u\varepsilon}(\tau) = \max(R_{u\varepsilon}(\tau)), \\ \text{else} \\ \rho_{u\varepsilon}(\tau) = \min(R_{u\varepsilon}(\tau)). \end{cases} \quad (19)$$

Then the validity tests for a properly identified model are derived as

$$\begin{cases} \rho_{\varepsilon\varepsilon}(\tau) = 1, & \tau = 0, \\ \rho_{\varepsilon\varepsilon}(\tau) = 0 & \text{otherwise}, \end{cases} \tag{20}$$

$$\rho_{u\varepsilon}(\tau) = 0 \quad \forall\tau. \tag{21}$$

Compared with the other correlation tests based methodologies, the new approach enhances the power of nonlinear model validity tests and significantly reduces the number of correlation plots. For large N the correlation function estimates given in (18) and (19) are still asymptotically normal with zero mean and finite variance in accordance with the central limit theorem (Bowker & Lieberman, 1972) and the standard deviations are $1/\sqrt{N}$ and the 95% confidence limits are therefore approximately $1.95/\sqrt{N}$.

To demonstrate the improved detection power of the new method with respect to the other previous methods, a simulation example was employed and shown in appendices.

4. Simulation studies

Unacceptable prediction errors will occur when an identified model structure is incorrect or the estimated parameters are biased. It is believed that the new correlation tests should be competently able to detect these problems. Two simulated examples were selected to demonstrate the theoretical insight and diagnostic procedure.

Example 1. Ill model structure.

Determining model structure including selection of function form, maximum nonlinear degree, maximum lag, and terms to be involved, is the first crucial part of any modelling procedure. An identified model cannot perform adequately when under-modelling occurs even if its parameter estimates are unbiased. The capability of the new correlation tests for validating identified models with incorrect structures was demonstrated through a simulated example. Consider a discrete time nonlinear dynamical system formulated as below:

$$\left.\begin{aligned} z(n) &= \sin(u(n-1)\pi) + \cos(u(n-2)\pi), \\ y(n) &= z(n) + e(n), \end{aligned}\right\} \tag{22}$$

where $\mathbf{z}$ and $\mathbf{y}$ denoted the noise free output sequence and measured noise perturbed output sequence, respectively. The input sequence $\mathbf{u}$ was a uniformly distributed random data sequence with zero mean and amplitude range ± 1. The noise $\mathbf{e}$ was assigned as a normally distributed random sequence with zero mean and variance of 4×10^{-4}. All these sequences were sampled with 1000 data points. Fig. 3 shows the map of $z(n)$ versus $u(n-1)$ and $u(n-2)$.

In the present study, nonlinear polynomial models with different degree of nonlinearity were selected to fit the data. Initially, the degree of nonlinearity was set to 3, and the maximum lag of the input was set to 2. By least squares

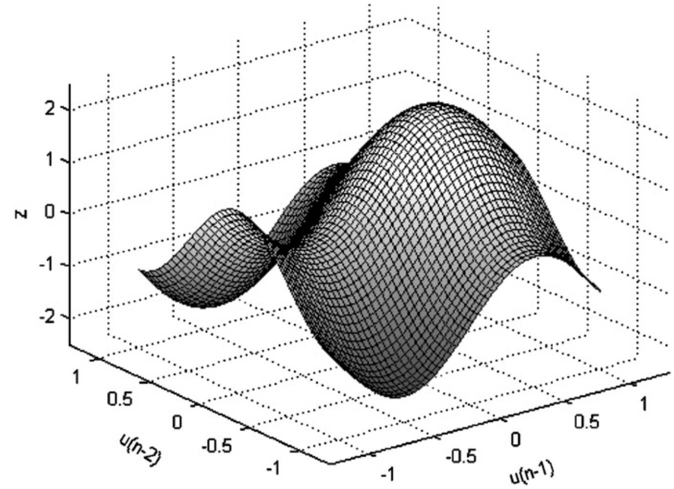


Fig. 3. Map of $z(n)$ versus $u(n-1)$ and $u(n-2)$ of model (22).

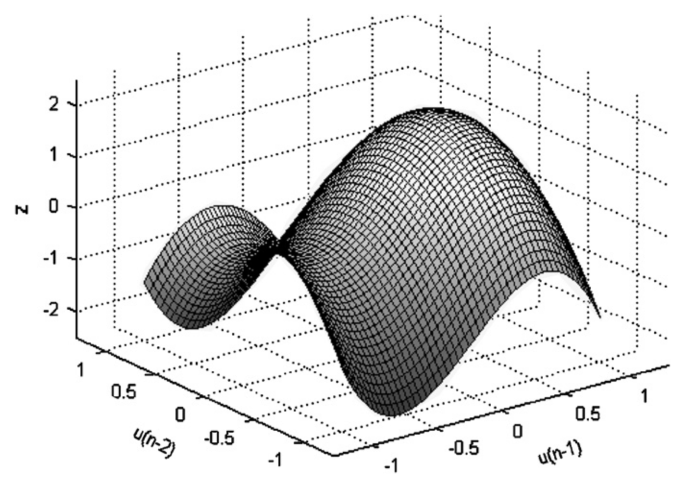


Fig. 4. Map of $\hat{z}_1(n)$ versus $u(n-1)$ and $u(n-2)$ of model (23).

parameter estimation, model was established as

$$\left.\begin{aligned} \hat{z}_1(n) &= 0.7384 + 2.6874u(n-1) + 0.012u(n-2) \\ &\quad -0.0155u^2(n-1) - 2.2328u^2(n-2) \\ &\quad -2.8765u^3(n-1) - 0.0505u^3(n-2), \\ \varepsilon_1(n) &= y(n) - \hat{z}_1(n), \end{aligned}\right\} \tag{23}$$

where $\hat{\mathbf{z}}_1$ and ϵ_1 denoted predictive output sequence and residual sequence, respectively. Fig. 4 shows the map of $\hat{z}_1(n)$ versus $u(n-1)$ and $u(n-2)$. It is evident that model (23) does not capture the underlying system characteristics at all. The surface in Fig. 4 is clearly different from the real system output versus input surface showed in Fig. 3.

Fig. 5 shows the results, obtained from using combined ODACF and combined ODCCF to test the residual ϵ_1. It is clear that $\rho_{u\varepsilon_1}(\tau)$ lies significantly outside the confidence interval at $\tau = 2$, which means the identified model (23) is invalid.

To improve the validity, the degree of nonlinearity of the polynomial model was increased to 4. Another trial model was

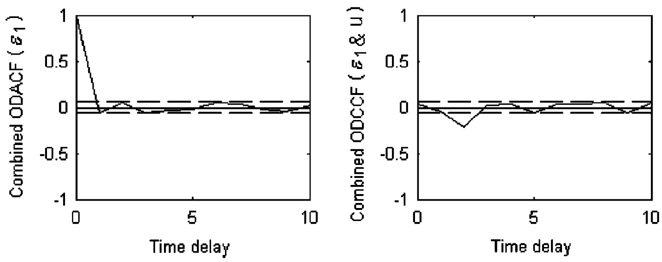


Fig. 5. Combined ODACF and combined ODCCF tests for model (23).

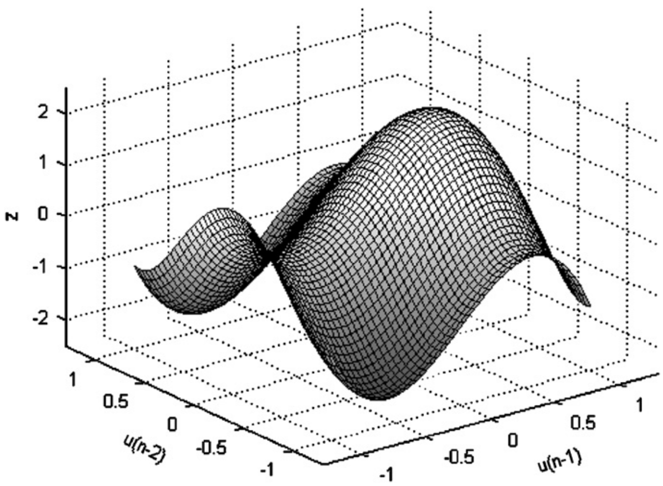


Fig. 6. Map of $\hat{z}_2(n)$ versus $u(n-1)$ and $u(n-2)$ of model (24).

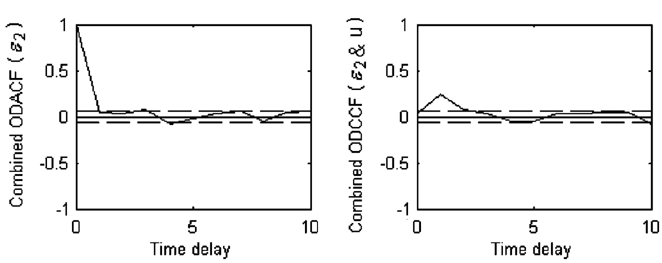


Fig. 7. Combined ODACF and combined ODCCF tests for model (24).

obtained

$$\left.\begin{aligned}\hat{z}_2(n) = {} & 0.9593 + 2.6881u(n-1) + 0.0092u(n-2) \\ & +0.0398u^2(n-1) - 4.4329u^2(n-2) \\ & -2.8771u^3(n-1) - 0.0116u^3(n-2) \\ & -0.0527u^4(n-1) + 2.5049u^4(n-2), \\ \varepsilon_2(n) = {} & y(n) - \hat{z}_2(n).\end{aligned}\right\} \quad (24)$$

Fig. 6 shows the map of $\hat{z}_2(n)$ versus $u(n-1)$ and $u(n-2)$, it can be observed that the mapping surface is improved compared to the first trial model, but still slightly different from Fig. 3.

Fig. 7 shows that model (24) is also invalid since the test of residual ϵ_2 by $\rho_{u\varepsilon_2}(\tau)$ lies significantly outside the confidence interval at $\tau = 1$.

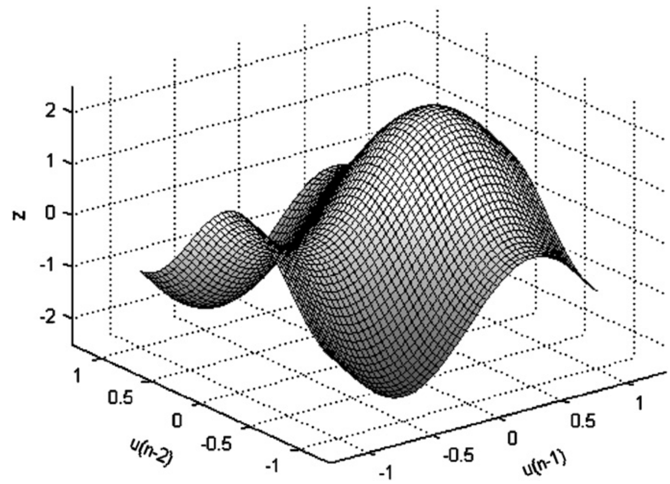


Fig. 8. Map of $\hat{z}_3(n)$ versus $u(n-1)$ and $u(n-2)$ of model (25).

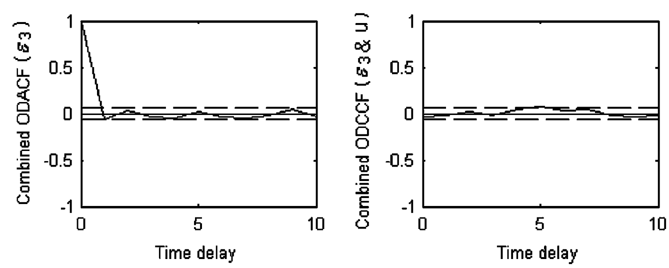


Fig. 9. Combined ODACF and combined ODCCF tests for model (25).

In fact, a polynomial model with the nonlinearity degree of 6 was determined to adequately approximate system (22). The system was identified as

$$\left.\begin{aligned}\hat{z}_3(n) = {} & 0.9882 + 3.1012u(n-1) + 0.0082u(n-2) \\ & +0.0217u^2(n-1) - 4.893u^2(n-2) \\ & -4.8035u^3(n-1) - 0.0316u^3(n-2) \\ & -0.0524u^4(n-1) + 3.8209u^4(n-2) \\ & +1.7155u^5(n-1) + 0.0273u^5(n-2) \\ & +0.0329u^6(n-1) - 0.9277u^6(n-2), \\ \varepsilon_3(n) = {} & y(n) - \hat{z}_3(n).\end{aligned}\right\} \quad (25)$$

Fig. 8 shows the map of $\hat{z}_3(n)$ versus $u(n-1)$ and $u(n-2)$. It can be observed that Figs. 3 and 8 display a high degree of similarity which suggests that model (25) is an adequate representation of system (22).

The results, obtained from using combined ODACF and combined ODCCF tests for residual ϵ_3, are illustrated in Fig. 9. Clearly they suggest that model (25) is a valid representation of (22).

Example 2. Biased parameters estimation.

Biased parameters estimation can induce invalid predictions despite the structure of an identified model being correct. To

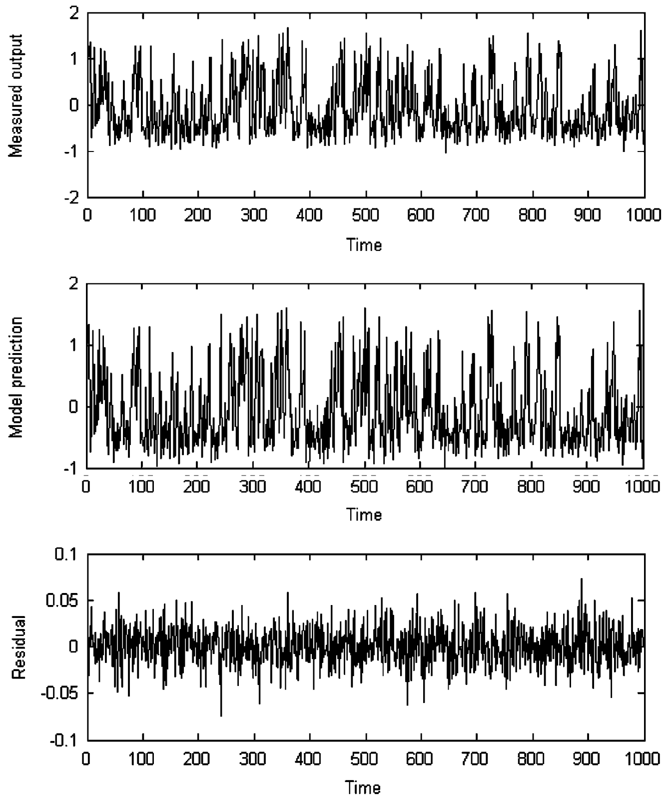


Fig. 10. The measured outputs, one step ahead model predictions, and the residuals of model (26).

illustrate the performance of the new model validity tests when the biased parameters estimation occurs, a nonlinear rational system was selected as below:

$$\left.\begin{aligned} z(n) &= (z(n-1) + u(n-1) + z(n-1)u(n-1) \\ &\quad + z(n-1)e(n-1))/(1 + z^2(n-1) \\ &\quad + u(n-1)e(n-1)), \\ y(n) &= z(n) + e(n). \end{aligned}\right\} \quad (26)$$

The category of system (26) often appears in the chemical and related fields (Dimitrov & Kamenski, 1991; Ford, Titterington, & Kitsos, 1989). In the present study, the parameters of all the terms in (26) were set to 1. Simulated output data were generated using system (26) with a uniformly distributed (zero mean and amplitude range from −1 to 1) random input sequence and a normally distributed (zero mean and variance of 4×10^{-4}) random noise sequence. A total of 1000 data samples were used in this case study. Fig. 10 shows the measured outputs, one step ahead model predictions, and the residuals respectively.

Fig. 11 shows the results obtained from using combined ODACF and combined ODCCF tests for system (26) that the identified model is valid. These results are obviously correct since the residual in model (26) is a completely random noise sequence.

To make a comparative study, consider a model estimated as (27) which has the same structure as system (26) but the

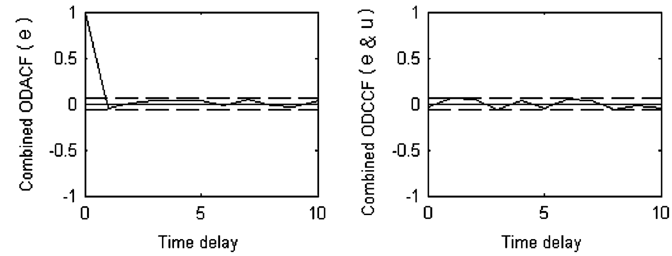


Fig. 11. Combined ODACF and combined ODCCF tests for model (26).

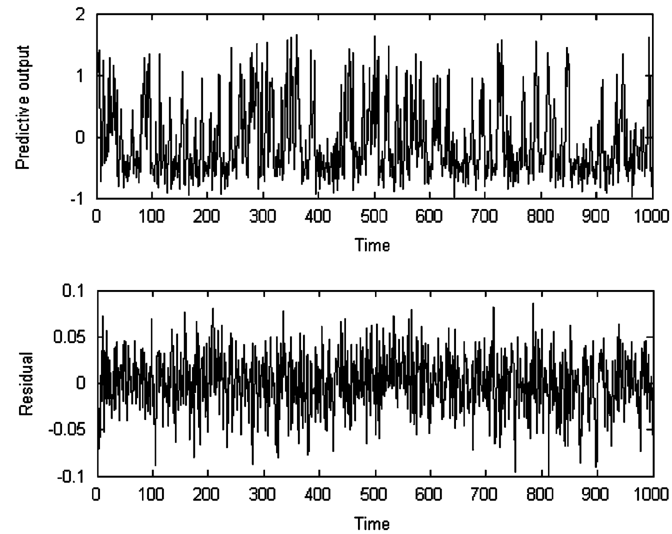


Fig. 12. The one step ahead model predictions, and the residuals of model (27).

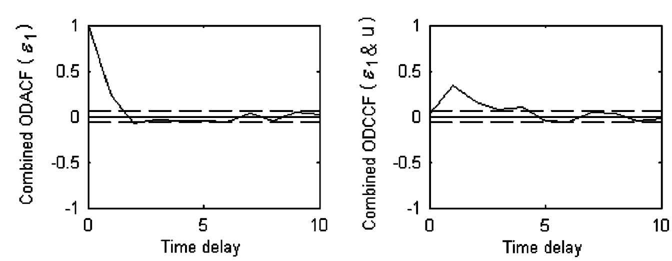


Fig. 13. Combined ODACF and combined ODCCF tests for model (27).

sole difference from model (27) is the estimated parameter associated with term $y(n-1)u(n-1)$.

$$\left.\begin{aligned} \hat{z}_1(n) &= (y(n-1) + u(n-1) \\ &\quad + 1.1y(n-1)u(n-1) + y(n-1)\varepsilon_1(n-1)) \\ &\quad /(1 + y^2(n-1) + u(n-1)\varepsilon_1(n-1)), \\ \varepsilon_1(n) &= y(n) - \hat{z}_1(n), \end{aligned}\right\} \quad (27)$$

where $\mathbf{y}$ denotes the same measured output sequence from system (26). Fig. 12 shows the predictive outputs and the residuals of model (27) and Fig. 13 shows the results obtained from using combined ODACF and combined ODCCF tests.

As illustrated, the combined ODACF test of residual ϵ_1 lies significantly outside the confidence interval at $\tau = 1$, and the

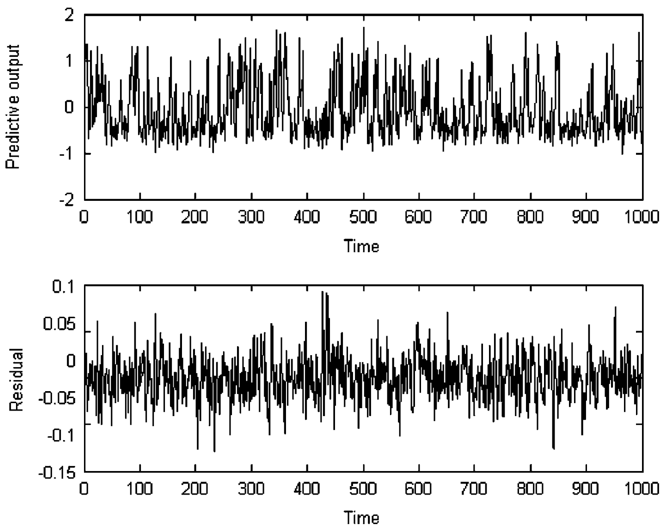


Fig. 14. The one step ahead model predictions, and the residuals of model (28).

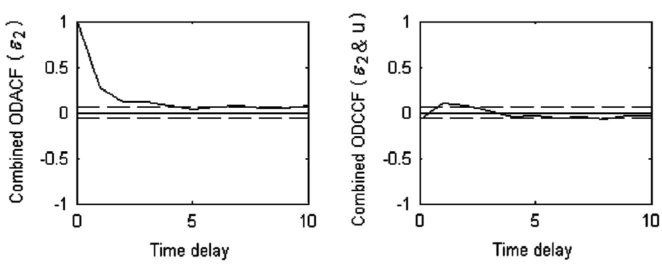


Fig. 15. Combined ODACF and combined ODCCF tests for model (28).

combined ODCCF test between input $\mathbf{u}$ and residual ϵ_1 also lies significantly outside the confidence interval at $\tau = 1$. Therefore, the estimated model is invalid to represent system (26).

Consider one more demonstration, a model with a biased parameter estimate associated with term $u(n-1)\varepsilon_2(n-1)$ was obtained as follows:

$$\left.\begin{aligned} \hat{z}_2(n) &= (y(n-1) + u(n-1) + y(n-1)u(n-1) \\ &\quad + y(n-1)\varepsilon_2(n-1))/(1 + y^2(n-1) \\ &\quad + 2.2u(n-1)\varepsilon_1(n-1)), \\ \varepsilon_2(n) &= y(n) - \hat{z}_2(n). \end{aligned}\right\} \quad (28)$$

Fig. 14 shows the predictive outputs and the residuals of model (28). Fig. 15 shows the results obtained from using combined ODACF and combined ODCCF tests that model (28) is clearly invalid.

5. Conclusions

A set of first order correlation tests have been proposed to weave a correlation detecting net for nonlinear model validity examination in the present study. The new procedure upgrades the detecting power of nonlinear associations compared with the other approaches that are based on higher order correlation functions. Two case demonstrations and one comparative study to the new method have been presented to illustrate that the new model validity tests appear to provide improved discriminatory performance and reduce number of correlation plots. It should be noted that in general an identified nonlinear model is only valid for the range of input excitation, if input range is changed, a new model need to be identified correspondingly, the developed validation procedure need to be applied to validate the model afterwards. It is believed that the new tests should be applicable to much wider class of nonlinear models including fuzzy systems and neural networks, which are under investigation currently and will be reported in following publications.

Acknowledgements

The authors are grateful to the anonymous reviewers for their helpful comments and constructive suggestions with regard to the revision of the paper.

Appendix A. Analytical proofs of ODCCFs

To analyse the feasibility of the new correlation functions dealing with each type of nonlinear association, three simple models illustrated in (5) that typically cover the four types of nonlinear associations are rewritten as below. It should be noticed that following examples are mainly used to demonstrate the ability to detect such correlationships in either noise free or additive Gaussian white noise situation. To reduce the derivation length, the noise free examples are analysed

$$y_1(n) = x_1(n)x_2(n), \quad (A.1)$$

$$y_2(n) = x_1^2(n), \quad (A.2)$$

$$y_3(n) = x_2(n)(x_1(n) + a). \quad (A.3)$$

In (A.3), a is a constant defined as

$$a \geqslant \max(|\mathbf{x}_1|). \quad (A.4)$$

For the convenience of the analysis, the ODCCFs expressed in (7)–(10) are further expanded in terms of mean removed (11) as follows.

Firstly, let

$$w_1 = \frac{1}{[(\sum_{n=1}^{N}(\alpha'(n))^2)(\sum_{n=1}^{N}(\beta'(n))^2)]^{1/2}}, \quad (A.5)$$

$$w_2 = \frac{1}{[(\sum_{n=1}^{N}(y'(n))^2)(\sum_{n=1}^{N}(\beta'(n))^2)]^{1/2}}, \quad (A.6)$$

$$w_3 = \frac{1}{[(\sum_{n=1}^{N}(y'(n))^2)(\sum_{n=1}^{N}(x'(n))^2)]^{1/2}}, \quad (A.7)$$

$$w_4 = \frac{1}{[(\sum_{n=1}^{N}(\alpha'(n))^2)(\sum_{n=1}^{N}(x'(n))^2)]^{1/2}}. \quad (A.8)$$

Since all the inputs are with zero mean, then (7)–(10) can be derived as follows:

$$r_{\beta'\alpha'}(\tau) = w_1 \sum_{n=\tau+1}^{N} \alpha'(n)\beta'(n-\tau)$$
$$= w_1 \sum_{n=\tau+1}^{N} (|y'(n)||x'(n-\tau)| - |y'(n)|\overline{|\mathbf{x}'|}$$
$$- \overline{|\mathbf{y}'|}|x'(n-\tau)| + \overline{|\mathbf{y}'|}\overline{|\mathbf{x}'|})$$
$$= (w_1 N)(\overline{|y'(n)||x'(n-\tau)|} - \overline{|\mathbf{y}'|}\overline{|\mathbf{x}'|}), \tag{A.9}$$

$$r_{\beta' y'}(\tau) = w_2 \sum_{n=\tau+1}^{N} y'(n)\beta'(n-\tau)$$
$$= w_2 \sum_{n=\tau+1}^{N} (y'(n)|x'(n-\tau)| - y'(n)\overline{|\mathbf{x}'|})$$
$$= (w_2 N)(\overline{y'(n)|x'(n-\tau)|}), \tag{A.10}$$

$$r_{x'y'}(\tau) = w_3 \sum_{n=\tau+1}^{N} y'(n)x'(n-\tau)$$
$$= (w_3 N)(\overline{y'(n)x'(n-\tau)}), \tag{A.11}$$

$$r_{x'\alpha'}(\tau) = w_4 \sum_{n=\tau+1}^{N} \alpha'(n)x'(n-\tau)$$
$$= w_4 \sum_{n=\tau+1}^{N} (|y'(n)|x'(n-\tau) - \overline{|\mathbf{y}'|}x'(n-\tau))$$
$$= (w_4 N)(\overline{|y'(n)|x'(n-\tau)}), \tag{A.12}$$

where overbar denotes the time average operation.

Proof. Using $r_{\beta_1'\alpha_1'}(0)$ to detect the first type of nonlinear association.

Consider the first model (A.1), the correlationships between $\mathbf{y}_1$ and both $\mathbf{x}_1$ and $\mathbf{x}_2$ exhibit the first type of nonlinear association and satisfy symmetrical properties 1 and 2. Since there is no delayed term included in the models and all the inputs are with zero mean, $\tau = 0$, $\overline{\mathbf{x}_1} = 0$, $\overline{\mathbf{x}_2} = 0$, $\mathbf{x}_1 = \mathbf{x}_1'$, $\mathbf{x}_2 = \mathbf{x}_2'$, $\overline{x_1(n)x_2(n)} = \overline{\mathbf{x}_1}\,\overline{\mathbf{x}_2} = 0$ and $\overline{(x_1(n)+1)x_2(n)} = (\overline{\mathbf{x}_1}+1)\overline{\mathbf{x}_2} = 0$.

Accordingly the ODCCFs between $\mathbf{x}_1$ and $\mathbf{y}_1$ can be derived as

$$r_{\beta_1'\alpha_1'}(0) = (w_1 N)(\overline{|x_1'(n)||y_1'(n)|} - \overline{|\mathbf{x}_1'|}\,\overline{|\mathbf{y}_1'|})$$
$$= (w_1 N)(\overline{|x_1(n)||x_1(n)x_2(n)|} - \overline{|\mathbf{x}_1|}\,\overline{|x_1(n)x_2(n)|})$$
$$= (w_1 N)(\overline{|\mathbf{x}_2|}(\overline{|\mathbf{x}_1|^2} - \overline{|\mathbf{x}_1|}^2)),$$

$$r_{\beta_1' y_1'}(0) = (w_2 N)(\overline{|x_1'(n)|y_1'(n)})$$
$$= (w_2 N)(\overline{|x_1(n)|(x_1(n)x_2(n))})$$
$$= (w_2 N)(\overline{|x_1(n)|x_1(n)}\,\overline{\mathbf{x}_2})$$
$$= 0,$$

$$r_{x_1'y_1'}(0) = (w_3 N)(\overline{x_1'(n)y_1'(n)})$$
$$= (w_3 N)(\overline{x_1(n)(x_1(n)x_2(n))})$$
$$= (w_3 N)(\overline{\mathbf{x}_1^2}\,\overline{\mathbf{x}_2})$$
$$= 0,$$

$$r_{x_1'\alpha_1'}(0) = (w_4 N)(\overline{x_1'(n)|y_1'(n)|})$$
$$= (w_4 N)(\overline{x_1(n)|x_1(n)x_2(n)|})$$
$$= (w_4 N)(\overline{\mathbf{x}_1|\mathbf{x}_1|}\,\overline{|\mathbf{x}_2|})$$
$$= 0.$$

Only $r_{\beta_1'\alpha_1'}(0)$ is a nonzero numeric that it can be used to detect the nonlinear association between $\mathbf{y}_1$ and $\mathbf{x}_1$. The ODCCFs for $\mathbf{x}_2$ can be proved similarly as for $\mathbf{x}_1$. □

Proof. Using $r_{\beta_1' y_2'}(0)$ to detect the second type of nonlinear association.

Consider the second model (A.2), the correlationship between $\mathbf{y}_2$ and $\mathbf{x}_1$ exhibits the second type of nonlinear association and satisfies the first symmetrical property. According to (A.9) to (A.12), the ODCCFs can be derived as follows:

$$r_{\beta_1'\alpha_2'}(0) = (w_1 N)(\overline{|x_1'(n)||y_2'(n)|} - \overline{|\mathbf{x}_1'|}\,\overline{|\mathbf{y}_2'|})$$
$$= (w_1 N)(\overline{|x_1(n)||x_1^2(n) - \overline{\mathbf{x}_1^2}|} - \overline{|\mathbf{x}_1|}\,\overline{|x_1^2(n) - \overline{\mathbf{x}_1^2}|})$$
$$= (w_1 N)(\overline{||x_1^3(n)| - |x_1(n)|\overline{\mathbf{x}_1^2}|}$$
$$- \overline{|\overline{|\mathbf{x}_1|}x_1^2(n) - \overline{|\mathbf{x}_1|}\,\overline{\mathbf{x}_1^2}|}),$$

$$r_{\beta_1' y_2'}(0) = (w_2 N)(\overline{|x_1'(n)|y_2'(n)})$$
$$= (w_2 N)(\overline{|x_1(n)|(x_1^2(n) - \overline{x_1^2(n)})})$$
$$= (w_2 N)(\overline{|x_1^3(n)| - |x_1(n)|\overline{\mathbf{x}_1^2}}),$$

$$r_{x_1'y_2'}(0) = (w_3 N)(\overline{x_1'(n)y_2'(n)})$$
$$= (w_3 N)(\overline{x_1(n)(x_1^2(n) - \overline{\mathbf{x}_1^2})})$$
$$= (w_3 N)(\overline{\mathbf{x}_1^3} - \overline{\mathbf{x}_1}\,\overline{\mathbf{x}_1^2})$$
$$= 0,$$

$$r_{x_1'\alpha_2'}(0) = (w_4 N)(\overline{x_1'(n)|y_2'(n)|})$$
$$= (w_4 N)(\overline{x_1(n)|x_1^2(n) - \overline{\mathbf{x}_1^2}|})$$
$$= (w_4 N)(\overline{x_a(n)(x_a^2(n) - \overline{\mathbf{x}_1^2})} + \overline{x_b(n)(x_b^2(n) - \overline{\mathbf{x}_1^2})}$$
$$+ \overline{x_c(n)(\overline{\mathbf{x}_1^2} - x_c^2(n))}),$$

where $\{\mathbf{x}_a | x_a^2(n) > \overline{\mathbf{x}_1^2}\}$, $\{\mathbf{x}_b | x_b^2(n) = \overline{\mathbf{x}_1^2}\}$, $\{\mathbf{x}_c | x_c^2(n) < \overline{\mathbf{x}_1^2}\}$ and $\{\mathbf{x}_a\} \cup \{\mathbf{x}_a\} \cup \{\mathbf{x}_a\} = \{\mathbf{x}_1\}$. Since $\{\mathbf{x}_1\}$ is a zero mean and uniformly distributed data sequence, $\{\mathbf{x}_1\}$, $\{\mathbf{x}_a\}$, $\{\mathbf{x}_b\}$ and $\{\mathbf{x}_c\}$ are symmetrical with $x=0$ and $\overline{\mathbf{x}_a}=\overline{\mathbf{x}_b}=\overline{\mathbf{x}_c}=0$. The function, therefore, can be expressed as

$$\begin{aligned}r_{x_1'\alpha_2'}(0) &= (w_4N)(\overline{x_a(n)(x_a^2(n)-\overline{\mathbf{x}_1^2})}+\overline{x_b(n)(x_b^2(n)-\overline{\mathbf{x}_1^2})}\\ &\quad+\overline{x_c(n)(\overline{\mathbf{x}_1^2}-x_c^2(n))})\\ &= (w_4N)(\overline{\mathbf{x}_a^3}-\overline{\mathbf{x}_a}\overline{\mathbf{x}_1^2}+\overline{0}-\overline{\mathbf{x}_c^3}+\overline{\mathbf{x}_c}\overline{\mathbf{x}_1^2})\\ &= 0.\end{aligned}$$

According to the above analysis, both $r_{\beta_1'\alpha_2'}(0)$ and $r_{\beta_1'y_2'}(0)$ are nonzero constants. The correlation value of $r_{\beta_1'y_2'}(0)$ is larger than that of $r_{\beta_1'\alpha_2'}(0)$ since both of $\overline{||x_1^3(n)|-|x_1(n)|\overline{\mathbf{x}_1^2}|}$ and $\overline{||\mathbf{x}_1|x_1^2(n)-\overline{|\mathbf{x}_1|}\overline{\mathbf{x}_1^2}||}$ are nonzero numbers. $r_{\beta_1'y_2'}(0)$, hence, can more effectively detect this type of nonlinear association. □

Proof. Using $r_{x_2'y_3'}(0)$ to detect the third type of nonlinear association and using $r_{x_1'\alpha_3'}(0)$ to detect the fourth type of nonlinear association.

Consider the third model (A.3), the correlationship between $\mathbf{y}_3$ and $\mathbf{x}_1$ exhibits the fourth type of nonlinear association and satisfies the second symmetrical property. Accordingly the ODCCFs for $\mathbf{x}_1$ can be derived as follows:

$$\begin{aligned}r_{\beta_1'\alpha_3'}(0) &= (w_1N)(\overline{|x_1'(n)||y_3'(n)|}-\overline{|\mathbf{x}_1'|}\overline{|\mathbf{y}_3'|})\\ &= (w_1N)(\overline{|x_1(n)||(x_1(n)+a)x_2(n)|}\\ &\quad-\overline{|\mathbf{x}_1|}\overline{|(x_1(n)+a)x_2(n)|}).\end{aligned}$$

Since $a \geqslant \max(|\mathbf{x}_1|)$, $|x_1(n)+a| = x_1(n)+a$ and the function can be derived as

$$\begin{aligned}r_{\beta_1'\alpha_3'}(0) &= (w_1N)(\overline{\mathbf{x}_1}\overline{|\mathbf{x}_1||\mathbf{x}_2|}+a\overline{|\mathbf{x}_1||\mathbf{x}_2|}\\ &\quad-a\overline{|\mathbf{x}_1||\mathbf{x}_2|}-\overline{\mathbf{x}_1}\overline{|\mathbf{x}_1||\mathbf{x}_2|})\\ &= 0,\end{aligned}$$

$$\begin{aligned}r_{\beta_1'y_3'}(0) &= (w_2N)(\overline{|x_1'(n)|y_3'(n)})\\ &= (w_2N)(\overline{|x_1(n)|(x_1(n)+a)x_2(n)})\\ &= (w_2N)(\overline{|x_1(n)|(x_1(n)+a)}\overline{\mathbf{x}_2})\\ &= 0,\end{aligned}$$

$$\begin{aligned}r_{x_1'y_3'}(0) &= (w_3N)(\overline{x_1'(n)y_3'(n)})\\ &= (w_3N)(\overline{x_1(n)(x_1(n)+a)x_2(n)})\\ &= (w_3N)(\overline{x_1(n)(x_1(n)+a)}\overline{\mathbf{x}_2})\\ &= 0,\end{aligned}$$

$$\begin{aligned}r_{x_1'\alpha_3'}(0) &= (w_4N)(\overline{x_1'(n)|y_3'(n)|})\\ &= (w_4N)(\overline{x_1(n)|(x_1(n)+a)x_2(n)|})\\ &= (w_4N)(\overline{\mathbf{x}_1^2}\overline{|\mathbf{x}_2|}).\end{aligned}$$

Therefore, $r_{x_1'\alpha_3'}(0)$ can properly detect the proper correlation between $\mathbf{x}_1$ and $\mathbf{y}_3$ which is the sole nonzero value.

The correlationship between $\mathbf{y}_3$ and $\mathbf{x}_2$ exhibits the third type of nonlinear association and the ODCCFs for $\mathbf{x}_2$ can be derived as follows:

$$\begin{aligned}r_{\beta_2'\alpha_3'}(0) &= (w_1N)(\overline{|x_2'(n)||y_3'(n)|}-\overline{|\mathbf{x}_2'|}\overline{|\mathbf{y}_3'|})\\ &= (w_1N)(\overline{|x_2(n)||(x_1(n)+a)x_2(n)|}\\ &\quad-\overline{|\mathbf{x}_2|}\overline{|(x_1(n)+a)x_2(n)|})\\ &= (w_1N)(a\overline{\mathbf{x}_2^2}-a\overline{|\mathbf{x}_2|}^2),\end{aligned}$$

$$\begin{aligned}r_{\beta_2y_3'}(0) &= (w_2N)(\overline{|x_2'(n)|y_3'(n)})\\ &= (w_2N)(\overline{|x_2(n)|(x_1(n)+a)x_2(n)})\\ &= (w_2N)(\overline{|x_2(n)|x_2(n)(x_1(n)+a)})\\ &= 0,\end{aligned}$$

$$\begin{aligned}r_{x_1'y_3'}(0) &= (w_3N)(\overline{x_1'(n)y_3'(n)})\\ &= (w_3N)(\overline{x_2(n)(x_1(n)+a)x_2(n)})\\ &= (w_3N)(a\overline{\mathbf{x}_2^2}),\end{aligned}$$

$$\begin{aligned}r_{x_1'\alpha_3'}(0) &= (w_4N)(\overline{x_1'(n)|y_3'(n)|})\\ &= (w_4N)(\overline{x_2(n)(x_1(n)+a)|x_2(n)|})\\ &= (w_4N)(\overline{\mathbf{x}_1\mathbf{x}_2|\mathbf{x}_2|}+a\overline{\mathbf{x}_2|\mathbf{x}_2|})\\ &= 0.\end{aligned}$$

Since $\overline{\mathbf{x}_2^2}$ and $\overline{|\mathbf{x}_2|}^2$ are nonzero and unequal values, $a(\overline{\mathbf{x}_2^2}-\overline{|\mathbf{x}_2|}^2) < a(\overline{\mathbf{x}_2^2})$. $r_{y_3'x_2'}(0)$ can more effectively detect the nonlinear correlation between $\mathbf{x}_2$ and $\mathbf{y}_3$ than $r_{\beta_2'\alpha_3'}(\tau)$. □

Appendix B. Numerical demonstration of the proved ODCCFs

To confirm the analytical proofs of the ODCCF tests, numerical demonstrations were performed to calculate ODCCFs for the three models. Inputs $\mathbf{x}_1$ and $\mathbf{x}_2$ were specified as uniformly distributed random data sequences with zero mean and amplitude range ±1. 1000 data points for each sequence were generated. The constant a in (A.3) and (A.4) was set as 1.

For model (A.1), Table B1 shows the simulation results obtained from using ODCCFs to accord with the analytical proofs. It is clear that $r_{\beta_1\alpha_1'}(0)$ can properly detect the first type of nonlinear associations.

Table B1
ODCCFs tests results for (A.1)

ODCCFs	$r_{\beta'\alpha_1'}(0)$	$r_{\beta'y_1'}(0)$	$r_{x'y_1'}(0)$	$r_{x'\alpha_1'}(0)$
$\mathbf{x}_1$	**0.6350**	−0.0461	−0.0494	0.0356
$\mathbf{x}_2$	**0.6480**	−0.0098	0.0518	−0.0542

Table B2
ODCCFs tests results for (A.2)

ODCCFs	$r_{\beta'\alpha_2'}(0)$	$r_{\beta' y_2'}(0)$	$r_{x' y_2'}(0)$	$r_{x'\alpha_2'}(0)$
$\mathbf{x}_1$	0.2584	**0.9685**	−0.0244	−0.0175

Table B3
ODCCFs tests results for (A.3)

ODCCFs	$r_{\beta'\alpha_3'}(0)$	$r_{\beta' y_3'}(0)$	$r_{x' y_3'}(0)$	$r_{x'\alpha_3'}(0)$
$\mathbf{x}_1$	−0.0315	−0.0476	−0.0359	**0.6482**
$\mathbf{x}_2$	0.6837	0.00238	**0.8767**	0.0015

For model (A.2), Table B2 shows the simulation results to accord with the analytical proofs, which $r_{\beta' y_2'}(0)$ is with the highest value that it can be used to effectively detect this type of nonlinear association.

According to the analytical proofs, $r_{x_1'\alpha_3'}(0)$ can properly detect the correlation between $\mathbf{x}_1$ and $\mathbf{y}_3$ and $r_{x_2' y_3'}(0)$ can more effectively detect the nonlinear correlation between $\mathbf{x}_2$ and $\mathbf{y}_3$. Table B3 shows the accordant simulated results for model (A.3).

Appendix C. A comparative study with an existing method

In this comparative study an example was selected to demonstrate the improved detection power and the reduction of the number of correlation plots of the new method with respect to Billings and Voon's nonlinear model validation method (Billings & Voon, 1986), that has been the most typical and popular one and widely applied in many nonlinear system identification applications during the last two decades. In Billings and Voon's method, it includes five first-order and higher-order correlation tests expressed as follows.

$$\begin{cases} r_{\varepsilon\varepsilon}(\tau) = \begin{cases} 1, & \tau = 0, \\ 0 & \text{otherwise}, \end{cases} \\ r_{u\varepsilon}(\tau) = 0 \quad \forall\tau, \\ r_{(u^2)'\varepsilon}(\tau) = 0 \quad \forall\tau, \\ r_{(u^2)'\varepsilon^2}(\tau) = 0 \quad \forall\tau, \\ r_{\varepsilon(\varepsilon u)'}(\tau) = 0 \quad \forall\tau, \end{cases} \tag{C.1}$$

where the dash $'$ denotes that the mean level has been removed from the corresponding data sequences.

Consider a rational system and three invalid residual sequences expressed as follows:

$$\left.\begin{aligned} z(n) = {} & (z(n-1) - 0.5z^2(n-1) \\ & + e(n-1)e(n-2) + u(n-1)u(n-2) \\ & + e^2(n-1))/(1 + z^2(n-1)), \\ y(n) = {} & z(n) + e(n), \end{aligned}\right\} \tag{C.2}$$

$$\begin{cases} \varepsilon(n) = 0.5y^2(n-1) + e(n), \\ \varepsilon(n) = 0.08y^2(n-1) + e(n), \\ \varepsilon(n) = (2e^2(n-1) + 3e(n-1)e(n-2)) \\ \qquad /(1 + e^2(n-1)) + e(n). \end{cases} \tag{C.3}$$

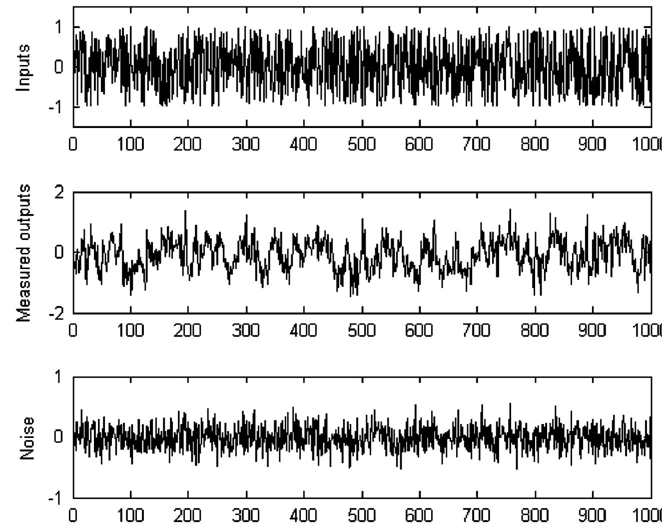


Fig. C1. The inputs, measured outputs and noise of system (C.2).

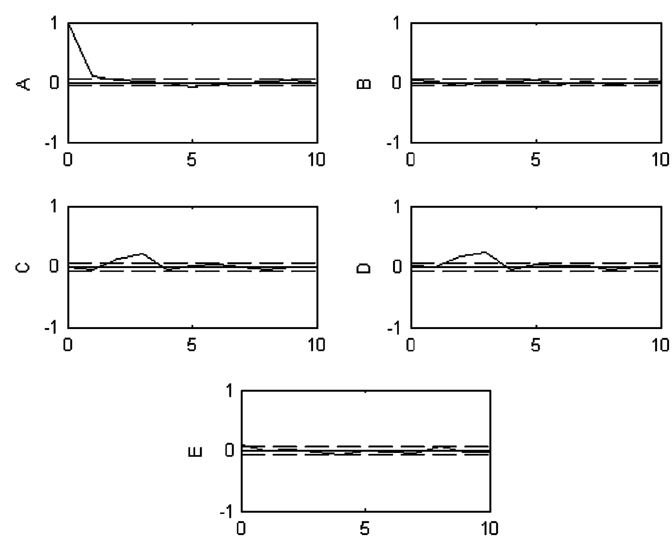


Fig. C2. Validation results obtained from using the Billings and Voon's tests for the first residual sequence in (C.3): A. $r_{\varepsilon\varepsilon}$; B. $r_{u\varepsilon}$; C. $r_{(u^2)'\varepsilon}$; D. $r_{(u^2)'(\varepsilon^2)'}$; E: $r_{\varepsilon(\varepsilon u)'}$.

The input sequence $\mathbf{u}$ was a uniformly distributed random data sequence with zero mean and amplitude range ± 1. The noise $\mathbf{e}$ was selected as a normally distributed random sequence with zero mean and variance of 4×10^{-2}. All these sequences were sampled with 1000 data points. Fig. C1 shows the inputs, measured outputs and noise of system (C.2). Figs. C2–C7 show the validation results for the three residual sequences in (C.3) obtained from using Billings and Voon's method and the new method. For the first residual sequence, $\varepsilon(n)$ is correlated to $y(n-1)$. In Figs. C2 and C3, $r_{(u^2)'\varepsilon}$, $r_{(u^2)'(\varepsilon^2)'}$ and $\rho_{u\varepsilon}(\tau)$ lie outside the confidence interval that both the two methods can be used to effectively detect this invalid residual sequence. For the second residual sequence, the effect from $y(n-1)$ in $\varepsilon(n)$ is reduced compared to the first residual sequence since the variance of $0.08y^2(n-1)$ is much smaller than that

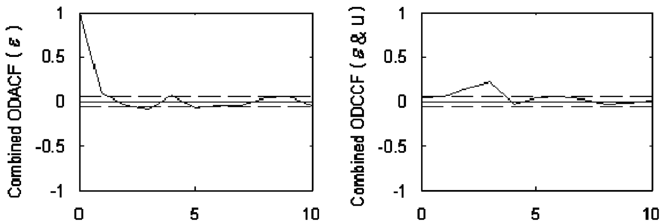


Fig. C3. Combined ODACF and combined ODCCF tests for the first residual sequence in (C.3).

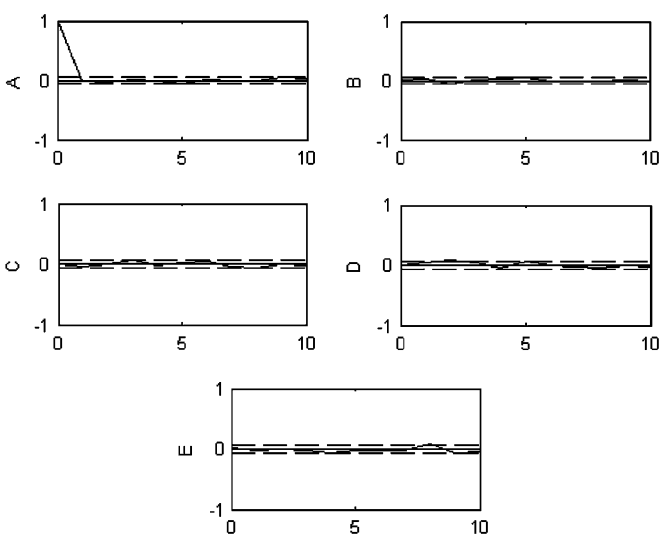


Fig. C4. Validation results obtained from using the Billings and Voon's tests for the second residual sequence in (C.3): A. $r_{\varepsilon\varepsilon}$; B. $r_{u\varepsilon}$; C. $r_{(u^2)'\varepsilon}$; D. $r_{(u^2)'(\varepsilon^2)'}$; E: $r_{\varepsilon(\varepsilon u)'}$.

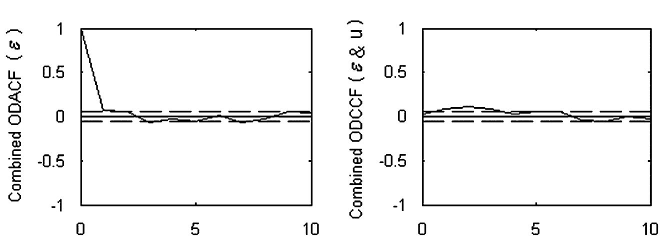


Fig. C5. Combined ODACF and combined ODCCF tests for the second sequence in (C.3).

of $0.5y^2(n-1)$. As shown in Fig. C4, Billings and Voon's method fails to detect the second residual sequence. Contrarily, in Fig. C5 $\rho_{u\varepsilon}(\tau)$ lies outside the confidence interval that the new method can be used to provide a more sensitive detection. Finally, consider the third residual sequence which is the extreme case and much more complex compared to the others. In Fig. C6, the Billings and Voon's method fails to detect this residual sequence. By using the new method, Fig. C7 clearly shows that $\rho_{\varepsilon\varepsilon}(\tau)$ lies outside the confidence interval at $\tau = 1$.

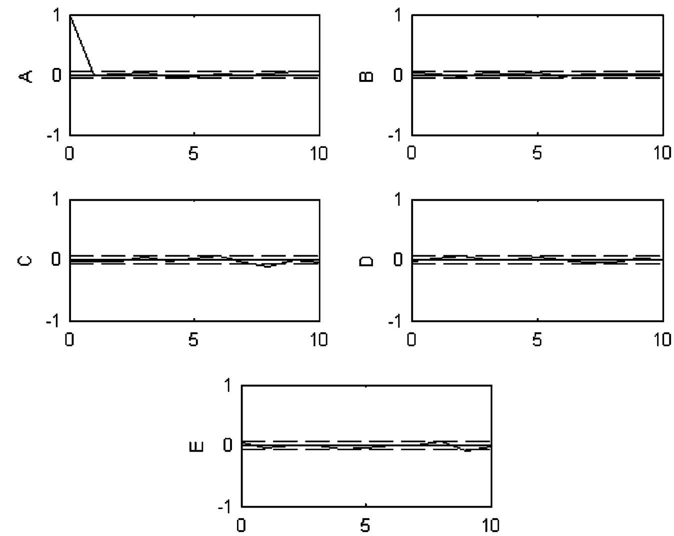


Fig. C6. Validation results obtained from using the Billings and Voon's tests for the third residual sequence in (C.3): A. $r_{\varepsilon\varepsilon}$; B. $r_{u\varepsilon}$; C. $r_{(u^2)'\varepsilon}$; D. $r_{(u^2)'(\varepsilon^2)'}$; E: $r_{\varepsilon(\varepsilon u)'}$.

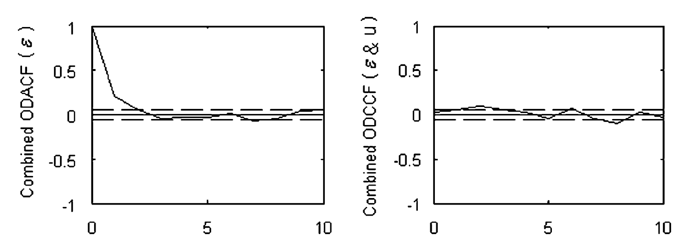


Fig. C7. Combined ODACF and combined ODCCF tests for the third residual sequence in (C.3).

References

Aguirre, L. A. (1995). A nonlinear correlation function for selecting the delay time in dynamical reconstructions. *Physics Letters A*, *203*, 88–94.

Aguirre, L. A. (1997). On the structure of nonlinear polynomial models: Higher order correlation functions, spectra, and term clusters. *Circuits and Systems I: Fundamental Theory and Applications*, *44*, 450–453.

Aguirre, L. A., & Billings, S. A. (1994). Validating identified nonlinear models with chaotic dynamics. *International Journal of Bifurcation and Chaos*, *4*, 109–125.

Billings, S. A., Chen, S., & Backhouse, R. J. (1989). The identification of linear and nonlinear models of turbocharged automotive diesel engine. *Mechanical Systems and Signal Processing*, *3*, 123–142.

Billings, S. A., & Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear system. *Proceeding of the Institution of Electronic Engineers, Pt D*, *130*, 193–199.

Billings, S. A., & Voon, W. S. F. (1986). Correlation based model validity tests for nonlinear models. *International Journal of Control*, *44*, 235–244.

Billings, S. A., & Zhu, Q. M. (1994). Nonlinear model validation using correlation tests. *International Journal of Control*, *60*, 1107–1120.

Billings, S. A., & Zhu, Q. M. (1995). Model validation tests for multivariable nonlinear models including neural networks. *International Journal of Control*, *62*, 749–766.

Bohlin, T. (1971). On the problem of ambiguities in maximum likelihood identification. *Automatica*, *7*, 199–200.

Bohlin, T. (1978). Maximum power validation of models without higher order fitting. *Automatica*, *14*, 137–146.

Bowker, A. H., & Lieberman, G. J. (1972). *Engineering statistics*. Upper Saddle River, NJ: Prentice-Hall.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis forecasting and control*. San Francisco: Holden-day.

Carson, E. R., & Flood, R. L. (1990). Model validation: Philosophy, methodology and examples. *Transactions of The Institute of Measurement and Control*, *12*, 178–185.

Dimitrov, S. D., & Kamenski, D. I. (1991). A parameter estimation method for rational functions. *Computers Chemical Engineering*, *15*, 657–662.

Dullerud, G., & Smith, R. (2002). A nonlinear functional approach to LFT model validation. *System and Control Letters*, *47*, 1–11.

Ford, I., Titterington, D. M., & Kitsos, C. P. (1989). Recent advances in nonlinear experimental design. *Technomatrics*, *30*, 49–60.

Hjalmarsson, H. (1993). Asymptotic correction tests in model validation. In *Proceedings of the 32nd conference on decision and control* (pp. 2058–2059). San Antonio, Texas, USA.

Ljung, L. (1999). *System identification theory for the user*. (2nd ed.), Upper Saddle River, NJ: Prentice-Hall.

Mao, K. Z., & Billings, S. A. (2000). Multi-directional model validity tests for non-linear system identification. *International Journal of Control*, *73*, 132–143.

Parke, J., Holford, N. H. G., & Charles, B. G. (1999). A procedure for generating bootstrap samples for the validation of nonlinear mixed-effects population models. *Computer Methods and Programs in Biomedicine*, *59*, 19–29.

Rangan, S., & Poolla, K. (1996). Time-domain validation for sampled-data uncertainty models. *IEEE Transaction on Automatic Control*, *41*, 980–991.

Soderstrom, T., & Stoica, P. (1990). On covariance function tests used in system identification. *Automatica*, *26*, 125–133.

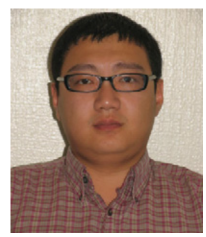
Sriniwas, G. R., Arkun, Y., Chien, I., & Ogunnaike, B. A. (1995). Nonlinear identification and control of a high-purity distillation column: A case study. *Journal of Process Control*, *5*, 149–162.

Thomson, M., Schooling, S. P., & Soufian, M. (1996). The practical application of a nonlinear identification methodology. *Control Engineering Practice*, *4*, 295–306.

Zhang, L. F., Zhu, Q. M., & Longden, A. (2007). A set of novel correlation tests for nonlinear system variables. *International Journal of Systems Science*, *38*(1), 47–60.

Zhu, Q. M., & Billings, S. A. (1997). Properties of higher order correlation function tests for nonlinear model validation. In *Proceedings of the 23rd international conference on industrial electronics, control and instrumentation* (vol. 1, pp. 306–310). New Orleans, USA.

Quan Min Zhu is Professor in control systems at the Faculty of Computing, Engineering and Mathematical Sciences (CEMS) University of the West of England (UWE), Bristol, UK. He had his higher education both in China and the UK, and obtained his PhD in Faculty of Engineering, University of Warwick, UK in 1989. His main research interest is in the area of nonlinear system modelling, identification, and control. Recently Professor Zhu started investigating electrodynamics of acupuncture points and sensory stimulation effects in human body, modelling of human meridian systems, and building up electro-acupuncture instruments. He has published over one hundred papers on these topics. Currently Professor Zhu is acting as the associate editor of International Journal of Systems Science (http://www.tandf.co.uk/journals/titles/00207721.asp) and editor of International Journal of Modelling, Identification and Control (http://www.inderscience.com/ijmic). Prof Zhu's brief CV can be found at http://www.ias.uwe.ac.uk/people.htm



Li Feng Zhang received the BSc degree in Electronic Engineering from Heilongjiang University, Harbin, China, in 1999, and MSc degree in RF and communication engineering from the University of Bradford, Bradford, UK, in 2003. He is currently enrolled in the PhD program at the Faculty of Computing, Engineering and Mathematical Sciences (CEMS), University of the West of England (UWE), Bristol, UK.

His research interests lie in the fields summarized as follows. Nonlinear dynamic system identification, mathematical modelling (e.g. NARMAX model), intelligence modelling (artificial neural networks (ANN) and fuzzy inference systems (FIS)), model validation, variable selection, and model structure selection.

Ashley Longden is a Senior Lecturer in control engineering in the Faculty of Computing, Engineering and Mathematical Sciences (CEMS) at The University of the West of England (UWE), Bristol, UK. After graduating from the University of Bath he worked on the research and development of digital controllers for aero engines for a period of ten years before joining UWE (formerly Bristol Polytechnic) in 1980.